METHODS FOR IDENTIFYING GENES ESSENTIAL TO THE GROWTH OF AN ORGANISM

FIELD OF THE INVENTION

The present invention relates to the use of high-density arrays or grids of genomic (or cDNA) libraries for the identification, sequencing and characterization of genes which are essential to the growth of an organism, and more specifically to a pathogen. The determination of these essential genes and the proteins encoded thereby is useful in the development of new therapies against such pathogens.

BACKGROUND OF THE INVENTION

Identification, sequencing and characterization of genes is a major goal of modern scientific research. By identifying genes, determining their sequences and characterizing their biological function, it is possible to employ recombinant technology to produce large quantities of valuable gene products, e.g. proteins and peptides. Additionally, knowledge of gene sequences can provide a key to diagnosis, prognosis and treatment in a variety of infectious diseases and disease states in plants and animals which are characterized by inappropriate expression and/or repression of selected genes or by the influence of external factors, e.g., carcinogens or teratogens, on gene function.

Methods have been described for the identification of certain novel gene sequences, referred to as Expressed Sequence Tags (EST). Adams *et al.*, *Science*, 1991, *252*:1651-1656. A variety of techniques have also been described for identifying particular gene sequences on the basis of their gene products. For example, see International Patent Application No. WO91/07087, published May 30, 1991. In addition, methods have been described for the amplification of desired sequences. For example, see International Patent Application No. WO91/17271, published November 14, 1991.

Genes which are essential for the growth of an organism, however, have been difficult to identify in such a manner as to be easily recovered for future analysis. The most common methodology currently employed to identify essential genes is a multi-

step process involving the generation of a conditionally lethal mutant library followed by the screening of duplicate members under the appropriate permissive and non-permissive conditions. Candidate mutants are then transformed with a second, genomic library and the desired genes isolated by complementation of the mutant phenotype. The complementing plasmid is recovered, subcloned, and then retested. However, this procedure comprises multiple subcloning steps to identify and recover the desired genes thus making it both labor intensive and time consuming.

Accordingly, there exists a need for a more efficient method of identifying genes essential to the growth of an organism.

SUMMARY OF THE INVENTION

In one aspect, the invention provides a method of identifying a gene or genes which are essential to the growth of an organism through the use of high density arrays or grids of genomic libraries. The method involves preparing a genomic library of a selected organism and providing a plurality of identical grids, each grid comprising a surface on which is immobilized at predefined regions on said surface a plurality of defined materials derived from the genomic library. The selected organism is then mutagenized, preferably by insertional mutagenesis, and grown in a test culture under a selected set of defined conditions. A control culture comprising the non-mutagenized selected organism is also grown under the same set of defined conditions. Surviving cells from the cultures are harvested and DNA from harvested cells of the mutagenized organism (test culture) and RNA, or DNA, from harvested cells of the non-mutagenized organism (control culture) are extracted and isolated. Labeled polynucleotide probes from the isolated DNA of the test culture and labeled polynucleotide probes from the isolated RNA (or DNA) of the control culture are then generated and hybridized to identical grids to produce a test hybridization pattern and a control hybridization pattern, respectively. Hybridization patterns on the grids are then compared to identify genes essential for growth of the selected organism. Essentiality of the identified gene for growth of the selected organism is then confirmed.

The method of the present invention may further comprise growing additional test cultures comprising the mutagenized organism and control cultures comprising the non-mutagenized organism under different sets of defined conditions. Labeled probes from the isolated DNA and RNA from these additional cultures are generated in the same fashion as previously described to produce test and control hybridization patterns for cultures grown under the different sets of defined conditions. Genes essential to the growth of the selected organism are then identified by comparing the hybridization patterns generated by mutagenized and non-mutagenized organisms grown under each of the different sets of defined conditions.

An additional aspect of the invention provides an isolated gene which is essential to the growth of an organism and is identified by one of the above methods.

Yet another aspect of the invention is an isolated protein produced by expression of the gene sequence identified above. Such proteins are useful in the development of therapeutic and diagnostic compositions, or as targets for drug development.

Yet another aspect of the invention is to identify broad spectrum antibiotics or antifungals which inhibit the expression of these essential genes.

In a related aspect, the present invention provides a method to identify conditionally lethal mutant genes of a selected organism by complementation with a non-mutagenized genomic library of the same organism. The method involves preparing a genomic library in either an integration vector, or in an expression vector, and providing a grid comprising a surface on which is immobilized at predefined regions on said surface a plurality of defined materials derived from the genomic library. The selected organism is then mutagenized, preferably by chemically induced point mutations, and grown (in a test culture) under permissive and non-permissive conditions to identify mutagenized organisms that contain conditionally lethal mutant genes. Organisms that contain conditionally lethal mutant genes are transformed with the prepared (i.e., non-mutagenized) genomic library and the transformed organisms, or cells, are grown under the same non-permissive conditions used to identify mutagenized organisms that contain the conditionally lethal mutant genes. Surviving cells are harvested and DNA is extracted and isolated. Labeled polynucleotide probes

from the isolated DNA are then generated and hybridized to the grid to identify genes essential for growth of the selected organism.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill in the art from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

DETAILED DESCRIPTION OF THE INVENTION

The biochemical basis of many bacterial resistance mechanisms to antibiotics is now known. These mechanisms alone, or in concert, are responsible for the escalating problem of antibiotic resistance seen both in hospital and community acquired infection. The principle approach by researchers to overcome these problems has been to seek incremental improvements in existing drugs. Although these approaches contribute somewhat to the fight against infection by such resistant pathogens, new approaches are needed.

Methods have now been developed for identifying genes and gene products essential to the survival of an organism. Genes and gene products identified by these methods are useful as molecular targets for drug discovery. The methods of the present invention are useful in determining the effect of the total absence of a gene or gene product on the survival of an organism.

I. Definitions

Several words and phrases used throughout this specification are defined as follows:

As used herein, the term "gene" refers to the genomic nucleotide sequence from which a cDNA sequence is derived. The term gene classically refers to the genomic sequence, which upon processing, can produce different cDNAs, e.g., by

splicing events. However, for ease of reading, any full-length counterpart cDNA sequence will also be referred to by shorthand herein as gene.

By "gene product" it is meant any polypeptide sequence encoded by a gene. The term "genomic library" is meant to include, but is not limited to, plasmid libraries, PCR products from genomic libraries, cDNA libraries and known sequences. Methods for the construction of such libraries are well known by those skilled in the art. In a preferred embodiment of the present invention, a genomic library is constructed in a suicide vector. It is also preferred that the constructed library be adjusted to minimize the number of complete genes present in a single genomic insert to approximately one gene. Techniques for this adjustment are well known to the skilled artisan.

"Isolated" means altered "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living animal in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. For example, with respect to polynucleotides, the term isolated means that it is separated from the chromosome and cell in which it naturally occurs.

By "organism" it is meant any single cell organism. Preferably this includes, but is not limited to, bacterium (including both gram negative and gram positive species), viruses and lower eukaryotic cells such as fungi, yeast, molds and simple multicellular organisms. Preferably, the organism is a pathogen.

The term "pathogen" is defined herein as any organism which is capable of infecting an animal or plant and replicating its nucleic acid sequences in the cells or tissue of that animal or plant. Such a pathogen is generally associated with a disease condition in the infected animal or plant. Such pathogens may include, but are not limited to, viruses, which replicate intra- or extra-cellularly, or other organisms such as bacteria, fungi or molds, which generally infect tissues or the blood. Certain pathogens are known to exist in sequential and distinguishable stages of development, e.g., latent stages, infective stages, and stages which cause symptomatic diseases. In

these different states, the pathogen is anticipated to rely upon different genes as essential for survival or for pathogenicity.

As used herein, the term "solid support" refers to any known substrate which is useful for the immobilization of a plurality of defined materials derived from a genomic library by any available method to enable detectable hybridization of the immobilized polynucleotide sequences with other polynucleotides in the sample. Among a number of available solid supports, one desirable example is the supports described in International Patent Application No. WO91/07087, published May 30, 1991. Examples of other useful supports include, but are not limited to, nitrocellulose, nylon, glass, silica and Pall BIODYNE C. It is also anticipated that improvements yet to be made to conventional solid supports may also be employed in this invention.

The term "grid" means any generally two-dimensional structure on a solid support to which the defined materials of a genomic library are attached or immobilized.

As used herein, the term "predefined region" refers to a localized area on a surface of a solid support on which is immobilized one or multiple copies of a particular clone and which enables hybridization of that clone at the position, if hybridization of that clone to a sample polynucleotide occurs.

By "immobilized", it is meant to refer to the attachment of the genes to the solid support. Means of immobilization are known and conventional to those of skill in the art, and may depend on the type of support being used.

II. Compositions of the Invention

The present invention is based upon the use of high density arrays or grids of genomic libraries as a means for rapidly identifying genes essential for the growth of an organism.

A. Preparation of genomic libraries

For this analysis a random genomic library for the target organism is prepared. The genomic DNA is isolated using standard procedures for molecular biology such as

those disclosed by Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989. The genomic library is then constructed in accordance with procedures described by Fleischmann et al. *Science*, 1995, *269*:496-512. For the purposes of the present invention, a genomic library can comprise a plasmid library, PCR products from a genomic library, or known sequences. In one embodiment, a suicide vector is used for preparation of the genomic library. Examples of suicide vectors which may be used in the present invention are well known in the art. See, for example, Booker *et al. Lett. Appl. Microbiol.* 1995 *21*:292-297; Steinmeitz, M. and Richter, R. *Gene*, 1994, *142*:79-83; Yu et al. *J. Bacteriol.* 1994 *176*:3627-34; and Quandt, J. and Hynes, M.F. *Gene*, 1993, *127*:15-21. In a preferred embodiment, a suicide vector containing the broad host range erythromycin (Erm) gene can be prepared in a commercially available plasmid such as pBluescript (pBS; Stratagene, La Jolla, CA). The Erm gene is isolated as a Taq1 restriction fragment from the vector pE194 (Hourinouchi, S. and Weisbaum, B. J. Bacteriology 1982, 150:804-812). The Erm containing fragment is ligated directly into Nael digested, CIP-treated pBS and transformed into HB101 cells. Transformants are screened by PCR to determine the presence of the Erm gene. Using this vector, two Erm positive isolates were confirmed by sequence analysis and designated pJMErmA4 and pJMErmD2. For library construction, genomic inserts are placed into the unique SmaI site present in the polylinker region. It is also preferred that the constructed library be adjusted to minimize the number of complete genes present in a single genomic insert. Techniques for making this adjustment to the library are well known to those skilled in the art.

B. Preparation of Grid

A plurality of materials derived from the genomic library are gridded onto a surface of a solid support at predefined locations or regions, preferably at 6X coverage. By "plurality of materials derived from the genomic library" it is meant to include, but is not limited to, bacterium containing individual clones spotted onto and grown on a surface of the solid support at predefined locations or regions; or plasmid clones isolated from said library, PCR products derived from the inserts from the

plasmid clones, or oligonucleotides derived from sequencing of the plasmid clones, which are immobilized to the surface of the solid support at predefined locations or regions.

Numerous conventional methods are employed for immobilizing these materials to surfaces of a variety of solid supports. See, e.g., Affinity Techniques, Enzyme Purification: Part P, Methods in Enzymology, Vol. 34, ed. W.B. Jakoby, M. Wilcheck, Acad. Press, NY (1971); Immobilized Biochemicals and Affinity Chromatography, Advances in Experimental Medicine and Biology, Vol. 42, ed. R. Dunlap, Plenum Press, NY (1974); U.S. Patent 4,762,881; U.S. Patent No. 4,542,102; European Patent Publication No. 391,608 (October 10, 1990); or U.S. Patent No. 4,992,127 (November 21, 1989).

One desirable method for attaching these materials to a solid support is described in International Application No. PCT/US90/06607 (published May 30, 1991). Briefly, this method involves forming predefined regions on a surface of a solid support, where the predefined regions are capable of immobilizing the materials. The method makes use of binding substrates attached to the surface which enable selective activation of the predefined regions. Upon activation, these binding substances become capable of binding and immobilizing the materials derived from the genomic library.

Any of the known solid substrates suitable for binding nucleotide sequences at predefined regions on the surface thereof for hybridization and methods for attaching nucleotide sequences thereto may be employed by one of skill in the art according to the invention. Similarly, known conventional methods for making hybridization of the immobilized materials detectable, e.g., fluorescence, radioactivity, photoactivation, biotinylation, energy transfer, solid state circuitry, and the like may be used in this invention.

C. Preparation and Growth of Mutagenized Organism

The organism of interest is mutagenized by transfection with either a randomly integrating transposon or similar insertional or transposable elements of known

sequence (e.g., Tn, IS, phage Mu, Ty element) or with a constructed suicide vector and allowed to grow under a selected set of defined conditions.

III. The Methods of the Invention

A. Identification of Genes

The present invention employs the compositions described above in methods for identifying genes which are essential to the growth of an organism. These methods may be employed to detect such genes, regardless of the state of knowledge about the function of the gene.

In one embodiment, a gene or genes which are essential to the growth of a selected organism are identified through the use of two or more identical high density arrays or grids of genomic libraries prepared from the selected organism. For this analysis, at least two identical high density grids or arrays are prepared. Each grid is prepared from a random genomic library for a selected organism, preferably in a suicide vector. A plurality of defined materials derived from the genomic library are then gridded onto a solid support, preferably at 6X coverage. The insert size of this library is adjusted to minimize the number of complete genes that might be present in a single insert. In a preferred embodiment, the target insert size is one complete gene. For bacteria, the average length of a complete gene is approximately 1 kb.

The selected organism is mutagenized by transfection with either a randomly integrating transposon or similar insertional or transposable element of known sequence, such as Tn, IS, Ty element or phage Mu, or with the constructed suicide vector. The mutagenized selected organism is then cultured under a selected set of defined *in vitro* or *in vivo* conditions to produce a test culture. In addition, a non-mutagenized selected organism is also cultured under the same set of defined conditions to produce a control culture. By "defined conditions" it is meant, but is not limited to, standard *in vitro* culture conditions recognized as normal (i.e., non-pathogenic) for a selected organism, as well as *in vitro* conditions which reflect or mimic *in vivo* pathogenic settings (conditions) such as heat shock, auxotrophic, osmotic shock, antibiotic or drug selection/addition, varied carbon sources, and aerobic or anaerobic conditions, and *in vivo*, pathogenic conditions. Preferably, such

conditions are predetermined to allow maximum growth of the non-mutagenized organism. The surviving cells are then harvested. Harvesting can be performed during various growth stages of the cells to ascertain the essentiality of a particular gene during different stages of growth. For example, harvesting can be performed during early logarithmic growth, late logarithmic growth, stationary phase growth or late stationary growth. RNA (or DNA) is then extracted and isolated from the harvested non-mutagenized celis of the control culture, while DNA is extracted and isolated from the mutagenized cells of the test culture using standard methodologies well known to those skilled in the art.

RNA (or DNA) extracted from the non-mutagenized cells of the control culture and DNA extracted from the mutagenized cells of the test culture are then used to generate labeled probes. The extracted, isolated DNA of the test culture serves as templates in primer extension reactions using oligonucleotide primers directed against a transposon/integrated vector sequence and which extends into the neighboring (i.e., flanking) nucleic acid sequence of the (genomic) DNA. Such primers will vary depending upon the mutagenesis/vector system employed. For example, in one embodiment, where the libraries constructed in the pJMErmA4 or pJMErmD2 vectors are used for both gridding and mutagenesis, primers designated against sequences which flank the SmaI cloning site are used. Examples of such primers include, but are not limited to:

5'-AATTAACCCTCACTAAAGGGAACA-3' (SEQ ID NO:1);

5'-TGTTCCCTTTAGTGAGGGTTAATT-3' (SEQ ID NO:2);

5'-GTAATACGACTCACGGAGGGGCGA-3' (SEQ ID NO:3); and

5'-ACGCCCCTCCGTGAGTCGTATTAG-3' (SEQ ID NO:4).

The extension reactions are performed using detectably labeled, i.e. radio- or fluorescent dye-labeled or biotinylated, nucleotides and controlled so that the extension products average approximately 200 base pairs (bp) in length. A number of methods exist for generating the primer extension products. In one embodiment, primer extension reactions are performed under the following conditions: A sample containing 15 pmoles of appropriate primer or primers, 5 pmoles extracted DNA, 30 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM DTT, 0.1 mM dATP, 0.1 mM ^{32}P-dCTP,

0.1 mM dGTP, 0.1 mM dTTP, 0.25 mM ddATP, 0.25 mM ddCTP, 0.25 mM ddGTP and water to 135 µl is prepared. This sample is then incubated at 75°C for 15 minutes; 50°C for 30 minutes and 37°C for 15 minutes. Klenow polymerase (75 units in a total volume of 15 µl) is then added and the sample is incubated for 30 minutes at 37°C. EDTA to 20 mM is the added. The sample is then extracted 1 time each with phenol, chloroform and isoamyl alcohol, followed by a second extraction with chloroform and isoamyl alcohol. The product is then precipitated with ethanol.

When RNA (or DNA) from the non-mutagenized organism is used to generate the probes, isolated RNA (or DNA) is labeled according to standard methods using random primers, preferably hexamers, and reverse transcriptase. Such methods are routinely performed by those skilled in the art.

These labeled products are then used as hybridization probes against the identical high density grids. Labeled probes prepared from DNA extracted from mutagenized cells of the test culture are hybridized to one identical grid, while labeled probes from the RNA extracted from the non-mutagenized cells of the control culture are hybridized to a second identical grid. The generated test hybridization patterns and control hybridization patterns are then compared. Genes essential for the growth of the selected organism are identified by determining differences at the predefined regions of the grids between the test hybridization pattern and the control hybridization pattern grown under the selected set of defined conditions

Alternatively, additional test cultures comprising the mutagenized selected organism and control cultures comprising the non-mutagenized selected organism are grown under different sets of defined *in vitro* and *in vivo* conditions. Hybridization patterns for labeled polynucleotide probes prepared from DNA of the additional test cultures and RNA of the additional control cultures are then generated in accordance with procedures described herein. Genes essential to the growth of the organism are then identified by comparing the hybridization patterns of the test and control cultures for each set of defined conditions with each other. In one embodiment, genes essential to the growth of the organism will be those common to all of the hybridization patterns for all the cells. In another embodiment, genes essential for

growth of a selected organism will hybridize under one set of growth conditions and will not hybridize under a different set of growth conditions.

In another embodiment, a pool of conditionally lethal mutants of the organism can be generated and transformed with a second (genomic) library constructed in a transposon/integration based vector. Transformants are reselected under the original conditionally lethal conditions and the rescued, surviving isolates used for probe generation and hybridization analysis as described above. For example, a temperature sensitive (ts) mutant library is prepared according to standard procedures and screened under permissive vs. non-permissive conditions to identify conditionally lethal mutants. The identified conditionally lethal ts mutants are pooled and transformed with a second, genomic library constructed in a transposon/integration based vector containing both a conditional and a selectable marker system. Examples of vectors for this second library include, but are not limited to, pMAK705 (Bloomfield et al. *Mol. Microbiol.* 1991 *5*:1447-1457) and pG+host5 (Biswas et al. *J. Bacteriol.* 1993, 175:3628-3635). The resulting transformants are retested or grown under the original temperature selection for lethality/essentiality. Survivors represent isolates containing integrated vector plus complementing genomic sequences. DNA from these survivors is then isolated and probes are generated as described in the preceding paragraphs, whereby hybridizing clones identify essential genes of interest.

In yet another embodiment, a conditionally lethal mutant library is prepared according to standard procedures, is constructed in an expression vector, and transformed with a selectable, genomic library. The genomic library is constructed using standard molecular biology techniques such that expression of the inserted genomic DNA is under control of vector-located promoter sequences, and preferably contains selectable and conditional markers. Examples of vectors containing inducible promoter systems include, but are not limited to, pFL10 (Lopez de Felipe et al. *FEMS Microbiol. Lett.* 1994, *122*: 289-295) and pUB110 (Zyprian, E. and Matzura, H. *DNA* 1986, *5*:219-225). In this embodiment, temperature sensitive lethal mutants are screened under temperature sensitive selection and under induction conditions for the vector-located promoter sequences. Surviving isolates represent clones where

transcription of the exogenous plasmid insert complements the mutant phenotype. Probes are generated against the plasmid inserts and hybridized against the grids.

Essentiality of the gene to the organism is confirmed by inactivating the identified gene in the selected organism, preferably using a single gene disruption procedure such as a knock out experiment, and culturing the selected organism under the same defined conditions.

Clones identified by the methods of the instant invention can be used directly for sequence analysis and for knockout experiments to confirm their essentiality to the growth of the organism. Alternatively, a gene sequence from the identified clone can be subcloned into a suitable vector for knockout experiments as is common in the art. Sequence analysis is performed using standard methodologies well known to those skilled in the art. Initial sequencing may be performed using the M13 universal forward and universal reverse sequencing primers which flank the multiple cloning site of the vector. The resulting sequences are analyzed using conventional computer programs. Results of said analysis are used in determining the potential usefulness of the individual clones as antimicrobial targets.

For knockout experiments, plasmid DNA from the identified isolates is purified and transformed in a non-mutagenized organism using standard molecular biology techniques. The transformed cells are grown under antibiotic selection for the vector sequence. Surviving cells represent site-specific insertional events into genes which are not essential for growth since knockout of an essential gene would result in no viable transformants. DNA is isolated from the surviving cells and used as a template to generate probes in accordance with previously described procedures and the grids reprobed for analysis. Additional gene knockout experiments can be performed in accordance with procedures described by, for example, Guiterrez et al. *J. Bacteriol.* 1996 *178*:4166-4175. Gene knockout experiments thus provide information on the effect of the total absence of the gene product.

B. Other Methods of the Invention

As is obvious to one of skill in the art upon reading this disclosure, the compositions and methods of the invention may also be used for other similar

purposes. For example, in one embodiment, this method can be used to monitor the effect of potential drugs on essential gene expression, both in laboratories and during clinical trials with animals, especially humans. Because the method can be readily adapted by altering growth conditions or the stage at which the cells are harvested, it can essentially be employed to identify essential genes of any organism, at any stage of development, and under the influence of any factor which can affect gene expression.

IV. The Genes and Proteins Identified

Application of the compositions and methods of this invention as above described also provides other compositions, such as any isolated gene sequence which is essential to the growth of an organism. Another embodiment of this invention is any isolated pathogen gene sequence found to be essential to the survival of the pathogen in a host. Similarly, an embodiment of the invention is any gene sequence identified by the methods described therein.

These gene sequences may be employed in conventional methods to produce isolated proteins encoded thereby. To produce a protein of this invention, the DNA sequences of a desired gene invention or portions thereof identified by use of the methods of this invention are inserted into a suitable expression system. In a preferred embodiment, a recombinant molecule or vector is constructed in which the polynucleotide sequence encoding the protein is operably linked to a heterologous expression control sequence permitting expression of the human protein. Numerous types of appropriate expression vectors and host cell systems are known in the art for mammalian (including human), insect, yeast, fungal and bacterial expression.

The transfection of these vectors into appropriate host cells, whether mammalian, bacterial, fungal or insect, or into appropriate viruses, results in expression of the selected proteins. Suitable host cells, cell lines for transfection and viruses, as well as methods for construction and transfection of such host cells and viruses are well-known. Suitable methods for transfection, culture, amplification, screening and product production and purification are also known in the art.

In one embodiment, the essential genes and proteins encoded thereby which have been identified by this invention can be employed as diagnostic compositions

useful in the diagnosis of a disease or infection by conventional diagnostic assays. For example, a diagnostic reagent can be developed which detectably targets a gene sequence or protein of this invention in a biological sample of an animal. Such a reagent may be a complementary nucleotide sequence, an antibody (monoclonal, recombinant or polyclonal), or a chemically derived agonist or antagonist. Alternatively, the essential genes of this invention and proteins encoded thereby, fragments of the same, or complementary sequences thereto, may themselves be used as diagnostic reagents. These reagents may optionally be detectably labeled, for example, with a radioisotope or colorimetric enzyme. Selection of an appropriate diagnostic assay format and detection system is within the skill of the art and may readily be chosen without requiring additional explanation by resort to the wealth of art in the diagnostic area.

Additionally, genes and proteins identified according to this invention may be used therapeutically. For example, genes identified as essential in accordance with this method and proteins encoded thereby may serve as targets for the screening and development of natural or synthetic chemical compounds which have utility as therapeutic drugs for the treatment of disease states associated with the organism. As an example, a compound capable of binding to a protein encoded by an essential gene thus preventing its biological activity may be useful as a drug component preventing diseases or disorders resulting from the growth of a particular organism. Alternatively, compounds which inhibit expression of an essential gene are also believed to be useful therapeutically. In addition, compounds which enhance the expression of genes essential to the growth of an organism may also be used to promote the growth of a particular organism.

Conventional assays and techniques may be used for screening and development of such drugs. For example, a method for identifying compounds which specifically bind to or inhibit proteins encoded by these gene sequences can include simply the steps of contacting a selected protein or gene product with a test compound to permit binding of the test compound to the protein; and determining the amount of test compound, if any, which is bound to the protein. Such a method may involve the incubation of the test compound and the protein immobilized on a solid

support. Still other conventional methods of drug screening can involve employing a suitable computer program to determine compounds having similar or complementary structure to that of the gene product or portions thereof and screening those compounds for competitive binding to the protein. Identified compounds may be incorporated into an appropriate therapeutic formulation, alone or in combination with other active ingredients. Methods of formulating therapeutic compositions, as well as suitable pharmaceutical carriers, and the like are well known to those of skill in the art.

Accordingly, through use of such methods, the present invention is believed to provide compounds capable of interacting with these genes, or encoded proteins or fragments thereof, and either enhancing or decreasing the biological activity, as desired. Thus, these compounds are also encompassed by this invention.

Numerous modifications and variations of the present invention are included in the above-identified specification and are expected to be obvious to one of skill in the art. Such modifications and alterations to the compositions and processes of the present invention are believed to be encompassed in the scope of the claims appended hereto.